VIDEO VERIFICATION OF IDENTITY (VIVID)

TECHNICAL INFORMATION

TABLE OF CONTENTS

VIDEO VERIFICATION OF IDENTITY (VIVID)

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY (DARPA)

1 GENERAL INFORMATION

1.1 Purpose

The purpose of this document is to provide amplifying information on the VIVID technology program.

1.2 Summary of Important Dates

11 February 2003	Proposal Due Date
21 March 2003	Anticipated Evaluation Completion/Notification
21 June 2003	Contract Awards

2 INTRODUCTION

2.1 Background

A recent Defense Science Board (DSB) Precision Targeting Study (August 2001) identified two key operational constraints for precision targeting: a) limiting friendly losses and b) minimizing unintended effects. While the study suggested that the use of smaller, more precise weapons is a means to achieve this goal, it also recognized that there is a need for visual identification of targets to avoid unintended strikes. This need is most pronounced in real-time targeting scenarios, where the weapon targeting-delivery timeline is compressed. The VIVID program goal is to provide an automated means of moving target verification using airborne video data sources. VIVID will support the entire targeting-delivery process by concentrating on three technology areas, confirmatory identification, multiple target tracking, and checking the impact area for potential collateral damage.

Precision targeting capabilities in recent engagements have experienced shortfalls. Included were the limited abilities to strike mobile missiles and hardened facilities in Desert Storm; an inability to target mobile Surface-to-Air Missiles (SAM)s in Bosnia; a lack of targeting capability in Kosovo against tanks under trees, near houses, tractors, buses, non-combatants, and with extensive Camouflage, Concealment & Deception; and unintended strikes such as the Chinese Embassy in Belgrade. Although the rate of occurrence of unintended strikes is low, even one unintended strike can have enormous political implications in today's environment. With a desire to implement automated delivery of many smaller precision weapons in the battlefield of the future, it is necessary to develop an automated means to verify the intended target right up to the point of weapon impact. Modern weapons are precise but Global Positioning Systems with

Inertial Navigation System (GPS/INS) weapons are blind, laser-guided munitions put operators in harm's way and seekers can locate but cannot identify the target.

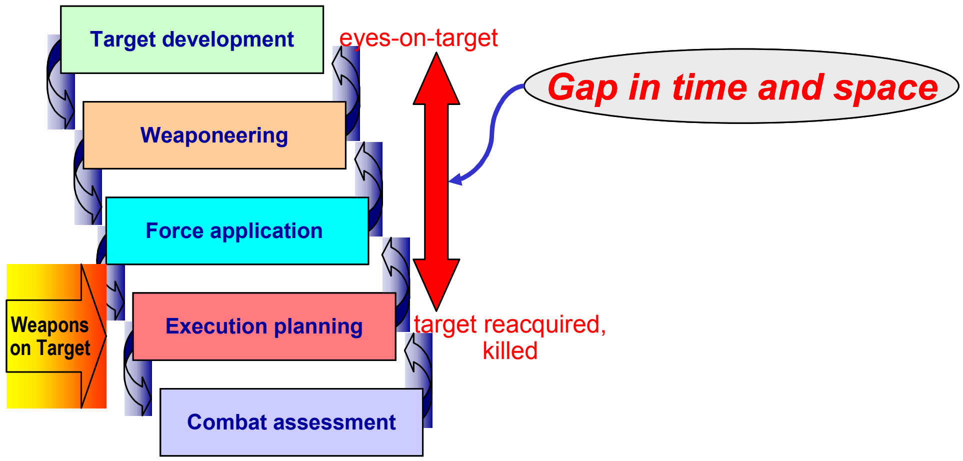

Figure 1:
Gap between target ID and weapon hit increases risk of collateral damage.

Current targeting methods introduce a gap in time and space between when the initial "eyes on target" occurs at target selection and when the final "eyes on target" occurs just prior to weapon delivery. This gap, shown in the figure above, results in targets being lost or modified with a resultant increase in the risk of collateral damage. The DSB Precision targeting study conducted in 2001concluded that "*older precision weapon targeting systems have become inadequate*".

There is a need to explore technological innovations using airborne platforms to find and designate targets according to a recommendation of the "Report to Congress - Kosovo/Operation Allied Force - After action Report - 31 Jan 01". VIVID will investigate technology to use precision identification to confirm the target identity to ensure that only the target is killed. GPS guided weapons have difficulty obtaining timely precision targeting information for moving targets. VIVID will attempt to develop technology to address this issue with moving target tracking from video, and automatic servo-control of a laser designator.

2.2 Objective

VIVID will focus on developing new technology to track, ID and strike moving targets using visible and infrared video data sources, and to do so under challenging conditions involving dense, urban traffic and frequent occlusions. This will require the handoff of the target from a broad-area surveillance cue, real-time geolocation and automated tracking to keep the sensor and a laser designator on target until weapon impact. Current generation video trackers lose track frequently when moving targets are temporarily occluded, shadowed, or in close proximity. VIVID will solve this problem by incorporating ID confirmation within the tracking loop. Manual control of a laser designator is currently not possible with round-trip SATCOM latencies of 1.5-2.0 seconds. VIVID will solve that problem by using its robust video tracker for servo-control of the laser-designator. VIVID will rely on existing technology for video geolocation and instead will focus on developing the robust extraction of moving targets from video and on the incorporation of automatic target verification to disambiguate confusers that arise from extended occlusions and/or dense traffic.

3 TECHNICAL APPROACH

3.1 System Considerations

Phase I of the VIVID Program will comprise one or more work packages in each of the technology areas listed below. Upon demonstrating the satisfactory achievement of Phase I performance goals, a distinct and separate procurement may be issued to select a VIVID System Integrator (SI) to assemble the selected technology components into an integrated Phase II VIVID system.

VIVID technology is envisioned to be embodied in software that can be run at a ground station or on the sensor platform. The software will be made interoperable with multiple platforms and sensors. The desired outcome is to track and target multiple moving targets from a single airborne sensor. A hardware interface is needed to control the sensor gimbal and zoom, and to control the laser-designator as well. No new communications channels will be needed and no hardware modification should be required. An increasingly complex set of demonstration criteria will be utilized to measure progress against program milestones. These criteria are detailed below in Section 5.

Because of the need to support strike operations both day and night, VIVID components must function with both visible and MWIR motion imagery. Solutions should not rely on the fusion of multiple video streams because small tactical UAVs do not usually carry more than one sensor. Similarly, solutions that rely on more than one platform are not preferred because of the added operational complexity.

Proposals for integration of VIVID technologies are not desired at this time. Integration will be performed in a second phase of VIVID after the first phase has demonstrated the ability to achieve the specified performance milestones for each of the component technologies.

3.2 Overview of VIVID technology components

Phase I of the VIVID program will focus on three technology areas and will conduct a series of critical technology demonstrations to select and apply the most promising automated target verification technologies. The program will conduct three parallel technology development efforts that focus on, 1) Confirmatory Identification, 2) Multiple Target Tracking, and 3) Collateral Damage Avoidance

1) Confirmatory Identification (CID) is performed to disambiguate confusers that arise during tracking, as well as to reconfirm the identity of the target

immediately prior to weapon impact. CID must be performed using data/models extracted from the initial observation of the target – prior target models cannot be assumed because of the potentially limitless variety of target classes and individual variations within each class. The challenge is to recover enough 3-D target information from video sources to allow it to be recognized when viewed later from a different perspective and illumination. Examples of on-the-fly modeling approaches include, but are not limited to, structure from motion, epipolar image analysis, and 2-D homographies.. Confirmatory ID could utilize a combination of approaches to confirm the presence or absence of targets and non-targets in dynamic scenes of interest. Example technical approaches include, but are not limited to, appearance matching, projective invariants, 4-D model matching, and the use of the fundamental matrix/trifocal tensor. The objective is to increase the probability of positive ID over time while minimizing the probability of false identification.

2) Multiple Target Tracking (MTT) controls the sensor gimbal and interprets the video frames to maintain track of up to 7 distinct vehicles simultaneously – especially when those vehicles are not all within the field of view at the same time. MTT will maintain track on multiple targets by slewing the sensor rapidly from target to target, and collecting just a few frames on each observation. Those frames can be used for updating the target state vector and continuing the track. The matching procedure embodied in the Confirmatory ID module can be used when necessary to disambiguate confusers that arise during tracking. Besides controlling the pointing direction of the gimbal, the MTT software may also control the degree of camera zoom. The key objective of MTT is to maintain track on targets even in the difficult conditions involving dense traffic and frequent occlusions. This track will also be used for real-time servo-control of a laser designator during the terminal phase of the weapon's flight.

3) Collateral Damage Avoidance (CDA) will search the vicinity of the predicted impact point for moving people and vehicles that are not intended as targets. Additionally, the avoidance of fixed structures that are included in a list of non-strike entities will also be enforced. The area that must be searched is that area from which any noncombatants could possibly come within range of the weapon during the last few seconds of flight before weapon at the time of impact, and is a function of the weapon effects radius, the motion of the target, and the maximum acceleration of vehicles and people in the given terrain. Collateral damage search will be conducted immediately prior to the last opportunity to divert the weapon. CDA must be performed with imagery that has the coarsest resolution possible, so as to permit search of a potentially large area with minimal or no sensor-gimbal motion.

The three technical areas are described in more detail in Section 4.

4 DESCRIPTION OF TECHNICAL AREAS

4.1 Confirmatory ID

4.1.1 Objective

A central component of the VIVID concept is the ability to confirm the identity of a target periodically while it is being tracked, and again immediately prior to weapon impact. The objective of this module is to confirm the identity of a target using data/models that were derived from observations of the target when the human operator nominated it. The target can be assumed to be a moving vehicle on the ground, although it will not be moving continuously and may start, stop, and maneuver frequently. It should not be assumed that the target is a military vehicle, as all manner of vehicles may present themselves as targets at one time or another. The observations will consist of a sequence of EO/IR video frames whose duration and resolution are consistent with that needed by humans for positive target ID. The model derived from those observations can be anything sufficient to allow the same target to be readily identified when viewed at a later time by the same type of sensor. The duration over which the identity of the target must be verified is up to 10 minutes. ID verification will happen periodically while the target is being tracked, although one cannot assume continuous observation of the target - occlusions, deep shadows, and the periodic gimbaling of the sensor to observe other targets will all cause gaps in coverage. These gaps will vary in length from 1 to 15 seconds. The objective of the Confirmatory ID module is 99% probability of correct association after a 10-second gap. This is to be obtained in a set of increasingly difficult operating conditions and with the minimum image resolution necessary. Periodically improving the target model through subsequent observations is permissible, and may be desirable to achieve maximum recall rate.

4.1.2 Background

Appearance-based matching has a rich history within the field of image understanding, and some of those representations and techniques may be relevant to VIVID Confirmatory ID. However, it should be clear that Confirmatory ID requires appearance models that are invariant to all the phenomena that can be expected while tracking a moving vehicle from a moving aircraft. (Note: See the Moving Target Tracking Section for details on the Vehicle to be tracked.) For example:

- Pose: As the aircraft moves and the vehicle turns, the vehicle will be observed from various elevation angles and from various headings.
- Illumination: Although the sun will not move appreciably during the 10-minutes or less that a vehicle is tracked, the angle between the sun and the video sensor line of sight may change significantly.
- Shadows: A moving vehicle is subject to moving in and out of shadows cast by buildings, trees, other vehicles, clouds, and even the observing aircraft.

- Articulations: Some vehicles have articulated parts, whose relative orientation may change while it is being observed. For example, towed trailers, missile launchers, and main gun tubes of tanks.
- Occlusion: The vehicle at times may be fully or partially obscured, by buildings, trees, other vehicles, clouds, or terrain features.

The object recognition community has used 3D models successfully to overcome some of these challenges. Extremely detailed models have been employed for automatic target recognition. In VIVID, we cannot assume that the target of interest has been modeled in advance - there are simply too many potential target types, and too many individual variations to expect that a valid model exists for more than a modest number of target types. Instead, the Confirmatory ID module must create its own model from the initial observations. That model need not be a complete description of the vehicle, nor need it even be three-dimensional, but it must be sufficient to allow unambiguous identification of that particular vehicle instance when viewed again at a later time.

Projective invariants are another technique for matching that has been studied extensively. Techniques for measuring and matching geometric invariants may also be relevant to VIVID Confirmatory ID, provided they can be extracted reliably and used to obtain sufficient discriminatory power to disambiguate potential confusers.

Many researchers have employed representations involving the Fundamental Matrix to relate information from multiple observations, and these too, may be relevant to Confirmatory ID.

There is a rich set of prior work that is potentially of value to achieve the VIVID objective for Confirmatory ID. It is up to the Offeror to select an approach that combines the techniques best suited for VIVID, to extend them as necessary, and to demonstrate that the performance objective can be achieved.

4.1.3 Key problems to be addressed

The Performance goal of Prob(association) >= 99% after a 10-second gap has been selected on the basis of systems analysis that showed this threshold to be necessary in order to achieve system-wide goals for target kill rates and collateral damage avoidance. As a result, it is an inflexible goal that must be obtained. Instead, other parameters can be varied as necessary to achieve that goal - for example, image resolution, aircraft altitude, and processor cost can be sacrificed as necessary to achieve it. The contractor will be expected to strike the right balance to obtain the optimum set of achievable performance characteristics.

4.1.4 Experimental support

Throughout the course of the VIVID Program, all developments will be tested extensively using representative imagery obtained from video sensors on airborne platforms. The Confirmatory ID contractor is required to develop and demonstrate his components using data supplied by the government. It will not be necessary to demonstrate this software in real-time during a flight - instead canned data will be employed.

The government will furnish video sequences observing one or more vehicles. Excerpts from those sequences will be excised to obtain video that is representative of what would be obtained by a sensor employing the simultaneous multiple-target tracking scheme described in the MTT section.

The data available with this document (see Section 7.0)BAA includes visible and MWIR sequences that can be used to demonstrate a Confirmatory ID capability. Offerors with existing capabilities to perform this task in whole or in part are encouraged to exercise their systems using the provided data and to submit those processing results along with the proposal. The BAA data is but a sampling of the more voluminous data to be provided as GFI to the successful Offeror.

4.2 Multiple Target Tracking (MTT)

4.2.1 Objective

With the advent of several new standoff precision weapons, such as the small diameter bomb, and the multiplicity of targets in many engagements, sensor assets, including the sensor, sensor platforms, and the crews that manage them, become a key limiting resource in precision strike. A central component of the VIVID concept is the ability to track multiple targets simultaneously with one sensor to improve the efficiency of use of sensors, sensor platforms, and UAV crews. By doing so, MTT enables a UAV to maintain continuous custody of potential targets while the targeting process is proceeding. Additionally, VIVID can significantly increase the kill rate for each weapon platform without requiring additional hardware or human resources.

The objective of this module is to develop a moving target tracking capability that can simultaneously and reliably track one or more targets from the point of weapon release through to the final seconds before weapon impact. The target can be assumed to be a moving vehicle on the ground, although it will not be moving continuously and may frequently start, stop, and maneuver. The target may or may not be a military vehicle, as all manner of vehicles may present themselves as targets at one time or another. Additionally, terrain, buildings, or other objects may temporarily occlude a target, and other vehicles at times may pass by the target or accompany it in close proximity.

One way a single sensor can follow multiple targets is by maintaining a track on each vehicle and rotating continuously through a target revisit cycle. The sensor is panned from one target to another, zoomed as necessary, allowed to dwell on a target long enough to update its track data, and then moved to the next target. The observations will consist of a sequence of video frames whose duration and resolution are consistent with that needed to confirm the re-acquisition of the target under track and to update track information. The information that the MTT derives for each track may include estimates of the target's motion, such as velocity and acceleration. Descriptions of the target's appearance and type may be generated independently, or MTT may take advantage of the Confirmatory ID (CID) module. The MTT initiates a new track each time it receives a target nomination along with the approximate location of the target and a target model formed by the CID module. The MTT must use some of its revisit cycle timeline to locate and re-confirm a new target's ID by calling the CID module and then create a track that becomes part of its revisit cycle. During the course of tracking an individual target, the MTT can call the CID module as necessary to reconfirm that the same target has been reacquired upon revisiting it. The MTT terminates a track when it is no longer able to re-acquire the target upon a revisit or when a weapon impacts the target or the target is handed off to a terminal-tracker and designator for terminal weapon guidance.

MTT approaches may take advantage of feature based tracking techniques exploiting the sorts of spatial and temporal information available in video EO and IR data. In addition, the MTT module can use the Confirmatory ID module as a target verification tool to increase its confidence that it has reacquired the correct target. Using the CID Module requires positioning the sensor on the target, zooming to obtain the required resolution if necessary, dwelling on the target, and calling the CID module to return a confirmation that the correct vehicle is in view along with a probability of correct ID. The CID module may be especially useful when the vehicle is in traffic or in proximity to confuser vehicles, or after a long gap in observation, such as that resulting from an extended occlusion. However, the CID module may require additional resolution and time and therefore lengthen the average revisit rate, potentially reducing the overall probability that multiple targets can be successfully tracked. Determining the correct situations and frequency of use of a CID module will be part of the MTT task.

The MTT module must meet several different objectives and must work with both EO and IR video data. These objectives may conflict with one another and require MTT contractors to develop approaches that make informed trades between them. As a minimum, the MTT module must track at least three [3] moving vehicles with one EO or IR sensor-gimbal configuration for the duration of 2 minutes with a 95% probability of maintaining a track on the original target throughout the two-minute period. These are minimum acceptable performance goals; general performance objectives to be addressed by developers of the MTT module include the following:

1. Track as many vehicles simultaneously with a single sensor as possible.
2. Track each target for the duration of a typical standoff precision-weapon's flight time from release to impact.
3. Keep the probability of maintaining track through the duration of weapon's flight time as high as possible.
4. Keep the probability of tracking the correct vehicle as high as possible: minimize the probability that a non-target vehicle will be mistaken for the target, for example upon a revisit, after an occlusion, or in traffic.
5. Develop moving target tracking techniques that can accommodate occasional revisit gaps of as much as 15 seconds. Approaches that can do so allow the multi-target tracking sensor to serve also as a target designator in the terminal phase of weapons delivery.

4.2.2 Background

There is a long history of research and development in target tracking, especially for tracking objects in the air from the ground and for tracking objects on or under the sea surface. Tracking ground vehicles from the air has received less although still considerable attention. DARPA has sponsored extensive research and development in the area of Ground Motion Target Indicator (GMTI) radar tracking of mobile ground targets. Some existing targeting pods on today's strike aircraft contain automatic trackers that can lock on a high contrast target as the aircraft moves and the target moves. However, existing video trackers can easily loose track as background-to-target contrast changes. Today's trackers can also not easily reacquire a target once it is occluded temporarily. The Video Surveillance and Monitoring (VSAM) and the Airborne Video Surveillance (AVS) programs funded limited investigations of advanced ground target tracking algorithms using EO and IR video data. Simultaneous tracking of multiple targets using either EO or IR sensors has yet to be addressed. Its solution requires innovative concepts encompassing the coordinated solution of several technical challenges in parallel. These challenges are numerous and include, but are not limited to, the following:

- Potentially long revisit-times required by a gimbaled video sensor with a small FOV that must be slewed between different targets in the FOR.
- Target variations in size and dynamics. Targets can be expected to range in size from a motorcycle to a large military vehicle. Dynamics can include realistic full maneuverability with speeds up to 90 MPH as well as Move-Stop-Move.
- Inaccuracies in pointing and re-pointing to a specific spot on the ground caused by uncertainties and inaccuracies in the control of and knowledge of the gimbal's degrees of freedom and the sensor platform's position and attitude.
- Imperfect or complete lack of knowledge of terrain elevation data and elevation variations due to man-made and natural terrain features further exacerbating the sensor pointing accuracy problem.
- Obscuration of targets by trees, buildings, smoke, etc., up to 25% of the time.
- Gaps in observation caused by target occlusions compounded with gaps caused by revisiting multiple targets. Gaps in observation can also result from in-flight maneuvering of the aircraft bearing the sensor and gimbal.
- Confusion of targets with other vehicles in the vicinity of the target, especially in the face of gaps in observation. A baseline scenario can be expected to have up to 12 vehicles in the FOR with excursions of dense traffic including hundreds of vehicles in the FOR.
- Widely varying and realistic differences in material construction that may adversely affect IR sensing.

- Various lighting and temperature conditions related to conducting operations anytime of the day or night as well as during various weather conditions.
- Changes in target appearance caused by changes in perspective and lighting as both the sensor and the target move and maneuver.
- Highly maneuverable ground vehicles create areas of location uncertainty that increase the difficulty of reacquiring the vehicle after a gap in observation.
- Conflicting resolution requirements: high resolutions provide richer data for feature-based tracking or CID while lower resolutions provide a larger FOV to reduce search time for reacquiring a target and therefore allow revisits that are more frequent or allow more targets to be tracked.

MTT developers can attack these problems with many different strategies. It is expected that successful MTT contractors will have a solid understanding of state of the art tracking approaches and propose innovative alternatives or improvements to them. In addition to outstanding single target tracking concepts, innovative approaches are needed to address simultaneous multi-target tracking. As a limited example of some possible approaches, MTT developers might consider some combination of non-uniform revisit strategies, multi-resolution schemes, varying dwell times, adjusting target revisit order based on track confidence, varying target revisit order and resolution based on minimizing sensor motion and settling time as well as based on track quality, higher order models of vehicle kinematics and dynamics, tracking non-target vehicles in the FOR as well as target vehicles, use of collateral data such as digital elevation and feature data or national reference imagery, and use of contextual information derived from the sensor data, such as extraction of roads and intersections if possible. Multi-sensor and sensor-platform coordination strategies could be used to increase overall performance by judicious handoffs of targets between sensors on the same or different platforms to reduce loss of track by minimizing occlusions. In addition, access to the CID module allows MTT module developers to consider strategies that trade higher revisit rates against higher confidence matches. These examples are representative and may or may not prove useful for MTT. Offerors are encouraged to consider additional and/or alternative approaches that will promise creative and innovative solutions to multi-target tracking.

4.2.3 Key problems to be addressed

A systems analysis based on the Predator WESCAM skyball sensor and gimbal characteristics has shown that the efficiency of a UAV and its crew can be increased by several hundred percent in terms of target kill rates for several scenarios with multiple target tracking if at least 3 targets can be tracked simultaneous for 2 minutes with a 95% probability of maintaining the track on the target. As a result, this is an inflexible goal that must be achieved. The contractor is strongly encouraged to propose and develop technical solutions that will significantly exceed these goals, especially techniques that increasing the performance well above 3 simultaneous targets while meeting the 95%

probability of maintaining track for 2 minutes. Improving track probability or lengthening the time of tracking while holding the other goals constant is also encouraged, but is less important than increasing the number of targets tracked simultaneously. Exceeding the minimum performance thresholds will enhance the likelihood that VIVID will result in technology that can be transitioned to operational use and that future opportunities will develop for MTT contractors.

4.2.4 Experimental support

Throughout the course of the VIVID Program, all developments will be tested extensively using representative imagery obtained from video sensors on airborne platforms. The Multiple Target Tracking contractor will be required to develop and demonstrate their components using data supplied by the government. The Government will furnish sensor, gimbal, and platform specifications and provide a sensor command and control interface for controlling the sensor's orientation and zoom and for obtaining platform and sensor position and orientation data.

The government will furnish video sequences observing multiple vehicles. Excerpts from those sequences will be excised to obtain video that is representative of what would be obtained by a sensor employing the simultaneous multiple-target tracking scheme described in this section. In years 1 and 2, canned data will be employed. The canned data will simulate multiple looks under the MTT module's control by sampling a spatial sub region from each frame that corresponds to the view that would have been obtained if the sensor were pointed to the position specified by the MTT module on each revisit.

The data furnished with this Technical Information document (see Section 7.0) includes visible and MWIR sequences that can be used to demonstrate a Moving Target Tracking capability. Offerors with existing capabilities to perform this task in whole or in part are encouraged to exercise their systems using the provided data and to submit those processing results along with the proposal. If the Offeror proposes the use of collateral data, such as map or image reference data, then the Offeror must address several additional issues in the proposal and during development. The collateral data used must correspond in quality, resolution, and accuracy with what is typically available for operations such as the recent ones in Kosovo and Afghanistan. In addition, the Offeror should address in their proposal the operational availability of such data, potential sources, and the risk to the approach and resulting performance impact if it is not available. MTT modules that take advantage of collateral data will be evaluated in terms of their performance both with and without using such collateral data. The Government will attempt to furnish or work with the contractor to secure required collateral data corresponding to the test data sets within reason and subject to its cost and availability.

If the Offeror proposes the use of contextual information derived from the scene, then they must address several additional issues in their proposal and during development. In the proposal the Offeror must describe their approach to deriving that contextual information automatically in enough detail to assess its likelihood of success and they must address both the additional risks entailed and expected performance of both the context extraction algorithms and the MTT modules performance in cases where context cannot be derived accurately. MTT contractors who proposed the use of contextual information must develop and demonstrate any proposed context extraction algorithms as part of their VIVID MTT effort and deliverable software suite. Their MTT module will be assessed on its performance both with and without using contextual information. Extraction of contextual information must fit into the processing timeline specified.

4.3 Collateral Damage Avoidance (CDA)

4.3.1 Objective

The VIVID program will develop techniques to avoid collateral damage. A key element of the VIVID process will be an automated collateral damage check during weapon delivery. This check will detect moving personnel and vehicles, and fixed no-strike objects in an area of interest and assess the potential for collateral damage effects. If the VIVID software determines that personnel, vehicles or other fixed structures are vulnerable to collateral damage, it can command the weapon to safe its fusing, alter its flight path or possibly engage an alternate target.

4.3.2 Background

Precision Weapon delivery with today's and future weapons technologies must be able to destroy any selected target without undue risk to friendly forces and without risking unwanted damage to non-targets (i.e. collateral damage). Although modern weapons are precise, certain weapon configurations, such as GPS/INS weapons, are blind and do not possess the means to "see the target" and inhibit collateral damage after weapon release. Laser guided weapons have lesser risk of collateral damage because an operator is maintaining a designator on the target up to weapon impact. This minimizes the risk of hitting the wrong target but puts the designator operator in harm's way. Seeker guided weapons have the ability to locate a target autonomously but currently don't have the ability to identify the target or target area well enough to preclude collateral damage.

In recent engagements such as Kosovo, the instances of collateral damage have been rare. ABC News reported that only 5 out of 1000 strikes hit noncombatant targets during Kosovo. The rate of collateral damage may be limited but the political cost of collateral damage can be enormous, as evidenced by the inadvertent strike on the Chinese Embassy in Belgrade, Yugoslavia, in 1999.

In his Congressional testimony on Kosovo lessons learned, Oct 1999, Lt.Gen. Short, (Commander, Allied Forces Southern Europe, and NATO's Air Component Commander) stated stated “Collateral damage drove us to an extraordinary degree. General Clark committed hours of his day dealing with the allies on issues of collateral damage. Restrictions that were placed upon us as a result of collateral damage incidents ... placed the crew at increased risk."

4.3.3 Key problems to be addressed

Procedures currently exist in the target development process to preclude strikes against no-hit targets. The VIVID process will continue to abide by the current no-hit target lists maintained and updated by DIA within the Modernized Integrated Database (MIDB). VIVID will add a level of automation to this manual no-hit target review process by ensuring that the target area is viewed and cleared for collateral damage up until the final seconds before weapon impact. There is a minimum time before impact (on the order of three seconds) when it will not be possible to alter the weapon aimpoint or command the weapon to disarm. VIVID will provide a positive safeguard against collateral damage for the terminal phase of weapon delivery up to this "lockout" period seconds before weapon impact.

The collateral damage checks will dynamically survey the target area as the weapon approaches target impact and assess the potential for unwanted personnel, vehicles or fixed targets to be within the blast radius of the weapon. Monitoring of the target impact area will be conducted by an EO or IR sensor on an unmanned air vehicle working in cooperation with the weapon or by the weapon sensor itself (if it is configured with a terminal sensor). The "mechanics" of this dynamic survey change as the weapon approaches the target. The area of search decreases and the number of moving objects that could enter the blast radius is reduced as the weapon gets closer and closer to the target. The dynamic collateral damage checking process must take this into account and must also incorporate a search method that "leads" moving objects so that areas the moving target is approaching are examined for potential no-hit targets. Application of collateral damage checks for moving targets will be a more challenging problem than for fixed targets.

The key technology to be developed is Video Moving Target Indication (MTI). The CDA component is to automatically search both visible and infrared motion imagery for moving persons, vehicles, or groups of persons. The image location of nontarget vehicles and persons visible within the video is to be determined.

4.3.4 Experimental support

Throughout the course of the VIVID Program, all developments will be tested extensively using representative imagery obtained from video sensors on airborne platforms. The Collateral Damage Avoidance contractor is required to develop and demonstrate his components using data supplied by the government. It will not be necessary to demonstrate this software in real-time during a flight - instead canned data will be employed.

The government will furnish airborne video sequences observing several ground targets. These sequences will span a wide range of environmental scenarios under various illumination conditions. The CDA component is to perform Video MTI to detect all persons and vehicles within the FOV.

The video data available with this document (see Section 7.0) includes sample visible and MWIR sequences in compressed (MPEG) format that can be used to demonstrate moving object detection. Offerors with existing capabilities to perform this task in whole or in part are encouraged to exercise their systems using the provided data and to submit those processing results along with the proposal. The data is but a sampling of the more voluminous data to be provided as GFI to the successful Offeror.

5 PERFORMANCE EVALUATION

AFRL Sensors Directorate COMPASE Center will periodically evaluate VIVID component performance to measure progress towards program goals. Target, background, and sensor conditions will be varied, and algorithm sensitivity will be measured. Testing and evaluation will occur using data that represents real-world conditions and that challenges the algorithms as per the technical descriptions above.

The technical performance goals for each component of VIVID processing are summarized in Table 1 Summary of VIVID Performance Goals. All performance goals are expected to be met using visible and infrared motion imagery individually.

- **Performance Thresholds** indicate the minimum performance required for each component. VIVID will not proceed to Phase II if the performance thresholds are not met.

Stretch Goals identify the more demanding performance desired for VIVID to realize its full potential. Evaluation shall be performed to determine whether the the stretch goals have been met. Contractors are expected to develop solutions that satisfy the strecth goals in their respective technical areas.

TECHNICAL AREA	PERFORMANCE THRESHOLDS	STRETCH GOAL
Collateral Damage Avoidance	Detect Moving People P_d=97.5% FAR = 0.5/km^2 @ 0.1 m GSD	Detect Moving People P_d=97.5% FAR = 0.5/km^2 @ 0.3 m GSD
Collateral Damage Avoidance	Detect Moving Vehicles P_d=97.5% FAR = 0.5/km^2 @ 0.6 m GSD	Detect Moving Vehicles P_d=97.5% FAR = 0.5/km^2 @ 1.2 m GSD
Confirmatory ID	95% after 10-second gap --on 5-way multiple choice test	99% after 10-second gap --on 5-way multiple choice test
Single-Target Tracking	90% after 2 minutes --up to 25% time occluded --up to 5 confuser vehicles	95% after 2 minutes --up to 25% time occluded --up to 5 confuser vehicles
Multiple-Target Tracking	3 vehicles --up to 25% time occluded --up to 5 confuser vehicles	7 vehicles --up to 25% time occluded --up to 5 confuser vehicles

Table 1 Summary of VIVID Performance Goals

5.1 CID and MTT Evaluation

The CID functions: CID algorithms will be evaluated on their ability to accurately recognize a previously observed target.

The MTT functions: MTT algorithms will be evaluated on their ability to accurately track designated targets for 2 minutes or more. Additional considerations are the number of vehicles accurately tracked and the length of track time. MTT is expected to meet both the single-target tracking and the multiple-target tracking performance goals.

Scenario: Vehicles will have realistic characteristics and range from motorcycle sized to military vehicle. They will be fully maneuverable and range in velocity from 0-90 MPH. A typical scenario will contain up to 12 vehicles in the FOR (up to 7 of which may be designated targets). Scenarios with dense traffic and several hundred vehicles will be tested as an excursion. In addition, vehicles may move through trees, buildings, smoke, etc. and be obscured up to 25% of the time. Due to operations during day or night under any weather conditions, lighting and temperature will vary greatly.

Data: Input data to the algorithms will be GPS coordinates of the platform, gimbal details to ascertain the FOR and FOV, and EO/IR video frames. CID will be given several frames from a target-acquisition phase to develop a model. After several models are built, CID will be given up to 4 frames of a potential target to determine if it is a specific target. After making the determination, CID will receive an input representing the tracker's confidence of it being that target, which can be used in updating its model if desired. MTT will be notified of a target and request gimbal and/or focal length changes as needed. Canned data of several vehicles not necessarily sensed on the same day or in the same area will be used to simulate the requested gimbal changes. If CID is requested, a stub program will reply with characteristics equivalent to the design goals. MTT output will be gimbal and focal length changes, CID requests, video frames of tracked targets and associated confidence level (for a coherent picture to an operator), and calculated GPS coordinates of targets.

CID Metrics/Goals: 99% of potential target requests are properly matched to the existing modeled targets. In addition, the false target identification will be measured to determine the number of times a non-target is incorrectly identified as a target.

MTT Metrics/Goals: 95% of objects designated to be tracked will be accurately (within 2 meters) tracked after 2 minutes. Additionally, the number of targets that can accurately be tracked will be measured.

5.2 Collateral Damage Avoidance (CDA) Evaluation

The CDA functions: CDA algorithms will be evaluated on their ability to monitor vehicle and personnel activity in the vicinity of target(s).

The CDA scenario: Humans and vehicles may be present in the target area. Varying amounts of CDA objects will be active around the target area. The human CDA object evaluation will consist of at least one adult, and the number may grow as large as a small crowd with random motion. The algorithms should detect motion at velocities greater than or equal to a slow adult gait. The vehicle CDA object evaluation will consist of at least one vehicle having realistic motion (stop-go-stop and constant change of direction). The size of the CDA vehicle will range from a motorcycle to a 2 x 4 meter military vehicle. The velocity of the vehicle will be greater than 0 MPH and less than 90 MPH. Position data on vehicles and personnel will not be available to the CDA algorithm.

The CDA data: Input data to the CDA algorithms will have a variable FOV and resolution to simulate UAV surveillance and the final weapon delivery phases (i.e. wide FOV initially with several minutes of monitoring followed by increasing resolution during the last minute of data). Human CDA algorithms will be evaluated with data of 0.1 meter GSD resolution, and the vehicle CDA algorithms will be evaluated with data at 0.6 meter GSD resolution. Algorithm input will be digital video clips of at least 15 seconds duration in uncompressed format with metadata (when available).). Vehicle and/or personnel activity will be occurring in the target area within the sensor FOV during some but not all video sequences. Development and test data will contain condition variations including CDA object size and velocity, and CDA number or density.

The CDA Metrics/Goals: The metric for CDA will be Probability of Detection (P_d) for humans and for vehicles. The performance threshold goal is to have a P_d of at least 97.5% for each. The other metric is the False Alarm Rate (FAR), with a threshold goal of a FAR of no more than 0.5/ km^2.

6 PROGRAM CONSIDERATIONS

6.1 Data Support and Delivery

The development and performance evaluation of airborne video analysis algorithms is challenging, due to the cost of performing airborne operations and the infinite variation of scenes, targets, lighting conditions, and viewing geometries. Innovative solutions to development and evaluation are desired to minimize the difficulty and cost while maximizing the realism. Offerors should clearly state their approach to development and evaluation. Because the performance goals form a go/no-go gate for Phase II, evaluation and verification of algorithm performance is critical to success in Phase I. Offerors' descriptions of their approach to measuring and verifying performance will be a key part of the proposal assessment process.

Sample airborne video data has been acquired and is currently available for Offerors (see Section 7). This data is only an example of the types of data that will be used during the VIVID program. Performers may use this data to demonstrate their capabilities as part of their proposal submission, but it is not required.

At the Program start, approximately 20 data clips, 10 infrared and 10 visible, will be made available to selected contractors. The data will be recorded and distributed in uncompressed format with metadata (where available).). A report of the data with descriptions of the video clips will be distributed that includes information details on platform/aircraft parameters (altitude, approximate slant range to target, duration of clip, data class). The data will be provided in digital format on TBD media.

During the course of the VIVID Program, there will be ongoing data collection to facilitate the technology development. The performers will have the ability to suggest video content and request specific scenarios. The performers do not have to use just the DARPA supplied video clips for their development efforts; however, all video clips used must be delivered to the Government and available for use by all performers. Developers will be evaluated against government-supplied video clips near the conclusion of Phase I of the program to impose a standard scenario for evaluating satisfaction of CID technology development goals.

If Performers generate or provide data, the cost of that data must be disclosed to the government, and it should be included as a deliverable to the government.

6.2 Sensor Hardware

While improvements in sensors and gimbals and their control will improve many aspects of VIVID performance, including moving target tracking, such improvements are outside the scope of VIVID. Various agencies are working in parallel on such developments under different programs. The Government will specify specific sensor, gimbal, and sensor-platform characteristics that MTT module developers must use. Instead, other parameters can be varied as necessary to achieve the performance goals - for example, image resolution (within the limits of the sensors array and optics), aircraft altitude, and processor cost can be sacrificed as necessary to achieve performance goals.performance goals. The contractor will be expected to strike the right balance to obtain the optimum set of achievable performance characteristics. It is up to the Offeror to select an approach that combines the techniques best suited for VIVID, to extend them as necessary, and to demonstrate that the performance objectives can be achieved given the sensor, gimbal, and platform characteristics as specified by the Government. These sensor-related specifications may evolve over the course of the VIVID program. If higher performance specifications are provided by the Government in the future, then the minimum performance goals will be adjusted to reflect the performance boost provided by those hardware and control improvements that are outside of VIVID's scope.

6.3 Program schedule

The VIVID Phase I schedule will be driven by progress towards the performance goals for each component. Nominally, we anticipate two-year efforts, with evaluation taking place periodically leading to satisfaction of Go/NoNo-go by month 18. However, offerors are expected to establish their own schedule for meeting their performance goals.

The schedule and cost for meeting the goals are included as criteria for proposal evaluation.

Semi-annual VIVID Team Meetings rotating among contractor sites will be held for the duration of Phase I. Government visits to contractor sites will occur between Team Meetings.

6.4 Multiple Proposals

Offerors wishing to bid on more than one element of VIVID may combine their bids into a single proposal. If bidding on more than one VIVID component, the SOW and the cost should be separated as options. Multiple proposals from a single organization or team are also permitted.

6.5 Innovation is desired

The information contained in this document is intended to provide background information on the VIVID concept as it exists at the time of this writing. Offerors are permitted, and in fact encouraged, to propose concepts and solutions that are superior to those contained herein, while generally satisfying the intent of the VIVID concept. Innovation in problem formulation, as well as problem solution, is desired.

7 Distribution of Experimental Video Data

Sample video data, in compressed format (MPEG) will be made available to Offerors on a CD-ROM for use in the proposal preparation process. This is not the actual data that will be provided for use in development and evaluation during the Phase I effort.. One CD-ROM will be provided to each organization upon request. To receive sample video data, email a request with SUBJECT LINE: “Video Data Request” to lhollan@schafercorp-ballston.com and include the following information in the email:

Name of Organization
Point of Contact (first and last name)
Surface mail address (No PO Boxes if possible)
Phone number where POC may be best reached
POC’s email address

8 Acronyms

AVS	Airborne Video Surveillance
BAA	Broad Agency Announcement
CDA	Collateral Damage Avoidance
CID	Confirmatory Identification
CPFF	Cost Plus Fixed Fee
CR	Cost Reimbursement
CRR	Central Contractor Registration

DARPA	Defense Advanced Research Projects Agency
DIRO	Directors Office (DARPA)
DO	Delivery Order
DSB	Defense Science Board
EO	Electro-Optical
FAQ	Frequently Asked Questions
FAR	Federal Acquisition Regulations
GMTI	of Ground Motion Target Indicator
GPS/INS	Global Positioning Systems with Internal Navigation Systems
GSD	ground sample distance
HBCU	Historically Black Colleges and Universities
IR	Infrared
ITAR	International Traffic in Arms Regulations
IXO	Information Exploitation Office
MI	Minority Institutions
MIDB	modernized integrated database
MTT	Multiple Target Tracking
MWIR	Mid-wave infrared
PIP	Proposer Information Pamphlet
SAM	Surface-to-Air Missiles
SI	System Integrator
SOW	Statement of Work
STT	Single Target Tracking
TBD	To be determined
Tfims	Technical-Financial Information Management System
UAV	Unmanned Air Vehicle
VIVID	Video Verification of Identity
VSAM	Video Surveillance and Monitoring